

4 Detecção de Silhueta

No decorrer deste capítulo é proposto um modelo de detecção da silhueta da mão capaz de lidar com os erros da segmentação e ruídos na sua morfologia. Num primeiro passo são considerados os filtros morfológicos, depois técnicas de detecção de bordas, e finalmente algoritmos de detecção, coleta e representação dos pontos que definem a silhueta da mão.

Na próxima seção é apresentada uma motivação geral da detecção de silhueta considerando-se a sua importância no plano deste trabalho. Em seguida é feito um breve resumo do modelo de detecção e finalmente apresenta-se em forma mais detalhada cada um dos seus passos, os resultados obtidos e algumas considerações finais.

4.1. Motivação

A capacidade de se reconhecer objetos em uma imagem depende muito da quantidade de informações que se conheça de cada objeto. Logo, extrair características dos objetos é uma etapa e tarefa fundamental para alcançar os objetivos no processo de reconhecimento.

A extração de características depende fortemente de como os objetos são representados computacionalmente. Por este motivo, é necessário um cuidado especial na escolha da representação dos objetos de tal maneira que o processo de detecção das características possa se dar da maneira mais natural possível. Para que o processo de reconhecimento possa ser realizado é necessário que os objetos de interesse sejam identificados e representados adequadamente.

Uma forma bastante comum de representação dos objetos e que foi utilizada neste trabalho é a representação baseada na silhueta. A silhueta é uma representação concisa e suficiente para capturar a morfologia do objeto.

Particularmente no caso da mão em movimento (objeto de estudo deste trabalho), muitas informações podem ser obtidas extraindo-se a silhueta ou contorno [16] a partir das imagens geradas pela segmentação de fundo. Se a cada momento pode-se conhecer a silhueta de mão, então pode-se saber as

diferentes formas que a mão está descrevendo ao longo do tempo. Isto ajuda muito e é uma grande vantagem, principalmente para objetos que mudam a sua morfologia ao longo do tempo. Interfaces baseadas em gestos, no caso da mão, poderiam facilmente ser suportadas com uma modelagem de silhueta.

A detecção de silhueta é o processo que determina quais são os pontos da imagem que fazem parte da silhueta. Neste processo, são coletados pontos de forma a capturar a morfologia do objeto e finalmente estruturá-los de modo que possam servir como dado de entrada para as etapas posteriores do reconhecimento.

No processo de detecção de silhueta são considerados três passos fundamentais. Em um primeiro passo, é feita uma abordagem de melhoria da segmentação; aqui são considerados filtros morfológicos para tentar corrigir e diminuir ao máximo os erros de segmentação e ruídos na morfologia do objeto. No segundo passo considera-se uma abordagem para a detecção das bordas dos objetos, que está diretamente influenciada pelo passo anterior. No último passo é feita a coleta e estruturação dos pontos que fazem parte da silhueta da mão. É fundamental que este passo seja capaz de lidar com os ruídos e erros ainda presentes após os dois passos anteriores.

Nas seções seguintes são apresentados, com mais detalhes, cada um dos passos do processo de detecção de silhueta.

4.2. Refinamento da Segmentação

O objetivo principal nesta etapa é corrigir ou pelo menos diminuir os erros e ruídos do processo de segmentação. Para isto, é utilizada uma abordagem de processamento de imagens baseada em filtros morfológicos que visam corrigir a imagem do objeto segmentado completando pequenos buracos e eliminando regiões isoladas de poucos pixels. Os filtros morfológicos utilizados foram: o filtro de Dilatação e o filtro de Erosão [34].

A Dilatação, em geral, faz com que os objetos se dilatam ou aumentem de tamanho, enquanto que a Erosão faz com que eles encolham. Ambos filtros atuam nas bordas internas e externas dos objetos. A quantidade e a forma como os objetos se dilatam ou encolhem depende fortemente da escolha de uma máscara. As máscaras mais comuns são a de vizinhança 4 (N_4) e de vizinhança 8 (N_8), que estão ilustradas na Figura 17.

A característica principal destes filtros é trabalhar com um ponto de referência (ponto do meio) e com as características da sua vizinhança. Assim, diferentes funções podem ser definidas nesta relação de vizinhança.

Considerando apenas imagens binárias (preto ou branco), um exemplo desta relação seria: atribuir a cor preta ao ponto de referência se todos os vizinhos, definidos na máscara, forem da cor preta também. Na Figura 17, além da geometria de vizinhança definida pela máscara, mostram-se os valores dos vizinhos, onde a cor preta representa o fundo e a cor branca representa os objetos.

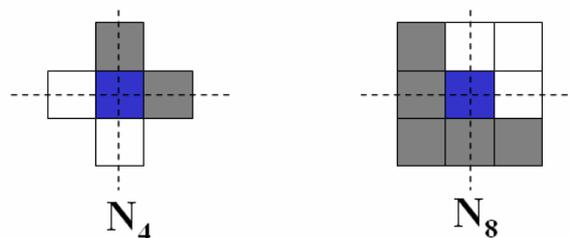


Figura 17 - Máscaras utilizadas nos filtros morfológicos de Dilatação e Erosão; vizinhança 4 (N_4) e vizinhança 8 (N_8) . Além da vizinhança definida pela máscara, mostram-se em preto os pontos que representam o fundo e em branco os pontos que representam os objetos.

Na aplicação dos filtros morfológicos nas imagens provenientes da segmentação, foi utilizada a máscara (N_8) . A seqüência de aplicação dos filtros foi: primeiro aplica-se a Dilatação e depois, na imagem dilatada, aplica-se o filtro de Erosão. Na Dilatação foi definido o seguinte critério: se um ponto faz parte do fundo e tem um vizinho que faz parte de um objeto, então ele também faz parte do objeto. Na Erosão, se um ponto faz parte do objeto e tem um vizinho que faz parte do fundo, então ele também faz parte do fundo.

Com a aplicação destes filtros na seqüência indicada, procura-se em princípio com a Dilatação expandir o objeto através das suas bordas internas e externas. Assim os buracos tendem a ser preenchidos e as bordas a serem expandidas uniformemente. Depois, aplicando a Erosão na imagem dilatada, procura-se retornar ao objeto original; apenas as bordas serão afetadas, e os buracos totalmente preenchidos na etapa de Dilatação serão mantidos. A aplicação destes filtros na seqüência indicada é conhecida também como

“fechamento”. Os resultados obtidos da aplicação desta seqüência de filtros morfológicos nas imagens de entrada são do tipo ilustrado na Figura 18.

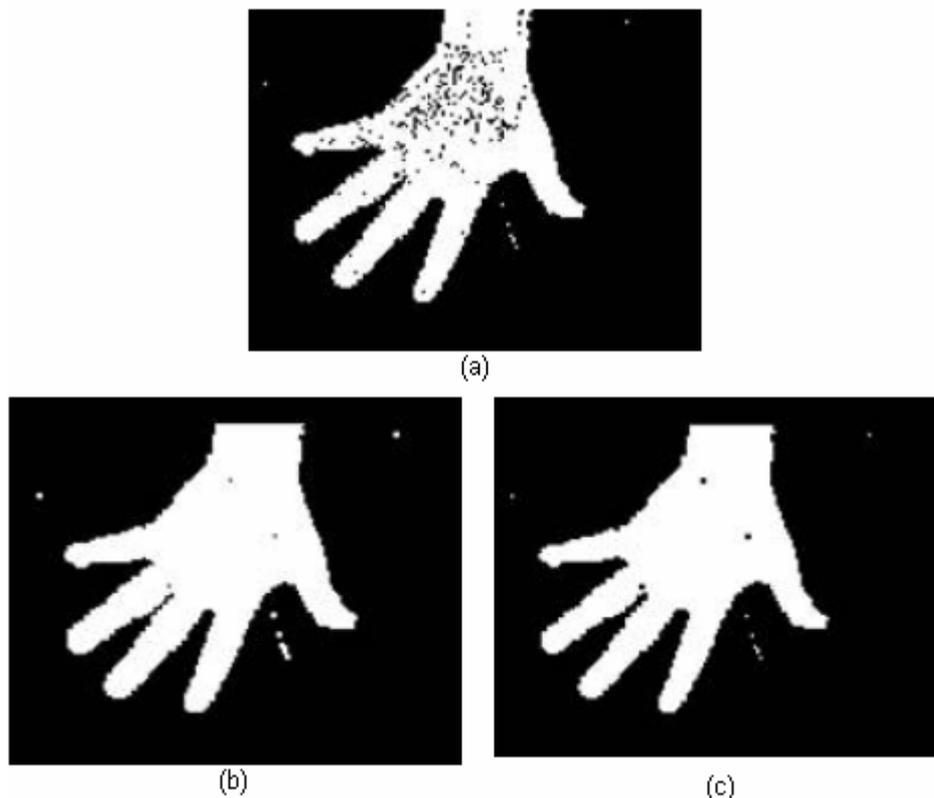


Figura 18 - Resultados obtidos da aplicação dos filtros morfológicos. Na seqüência (a) imagem resultado da segmentação; (b) resultado da aplicação da Dilatação em (a); e (c) resultado da aplicação da Erosão em (b). Considera-se (c) como resultado final desta etapa. Número médio de quadros por segundo igual a 53 qps.

4.3. Detecção de Bordas

O objetivo principal desta etapa é localizar os contornos internos e externos dos objetos, dado que estes contornos são a representação da silhueta. A conectividade do contorno é um fator importante; essa propriedade garante uma vantagem para os algoritmos no processo de detecção e coleta dos pontos que representam a silhueta da mão. Considera-se como a saída deste processo uma imagem binária onde os objetos são representados apenas pelos seus contornos.

Existem diversas abordagens para a obtenção de contornos em imagens binárias, como, por exemplo, as ilustradas por Trucco & Verri [38]. Tipicamente estas abordagens estão baseadas em medidas de gradiente ou na aplicação de filtros que exploram as características dos contornos. No presente trabalho, para

encontrar as bordas nesta etapa foi considerada uma abordagem diferente, que explora os resultados da aplicação dos filtros morfológicos nas imagens da segmentação. Na Dilatação as bordas dos objetos se expandem, e na Erosão essas mesmas bordas se contraem, então através da diferença destas imagens (a dilatada e a erodida) podemos achar uma aproximação das bordas dos objetos. Esta abordagem tem uma vantagem, pois aproveita os resultados da etapa anterior, tornando o processo computacionalmente barato comparado com os outros. Além disso, não incorpora medidas de limiar e depende somente do tamanho da máscara. Outra vantagem é que garante a conectividade dos contornos. Os resultados da aplicação deste processo com as imagens (b) e (c) da Figura 18 são mostrados na Figura 19.

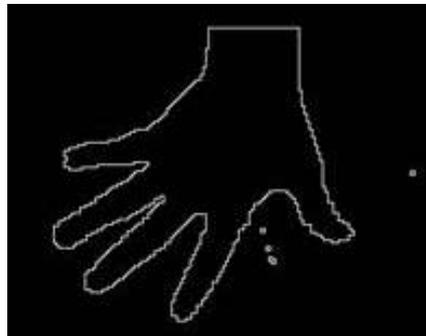


Figura 19 - Obtenção de contornos a partir da subtração das imagens (b) e (c) da Figura 18. O objeto “mão” é representado apenas pelo seu contorno, o mesmo acontece com o ruídos e erros da segmentação. Número médio de quadros por segundo igual a 45 qps.

A Figura 19 mostra o resultado para a mão totalmente estendida com a silhueta externa. Alguns testes foram feitos com diferentes posições da mão e a presença de contornos internos; esses resultados são mostrados na Figura 20. Nota-se nesta figura ainda alguns ruídos e erros da segmentação, os quais serão removidos na próxima etapa do processo de detecção de silhueta. As figuras em preto e branco são consideradas como sendo o resultado da etapa de detecção de contornos.

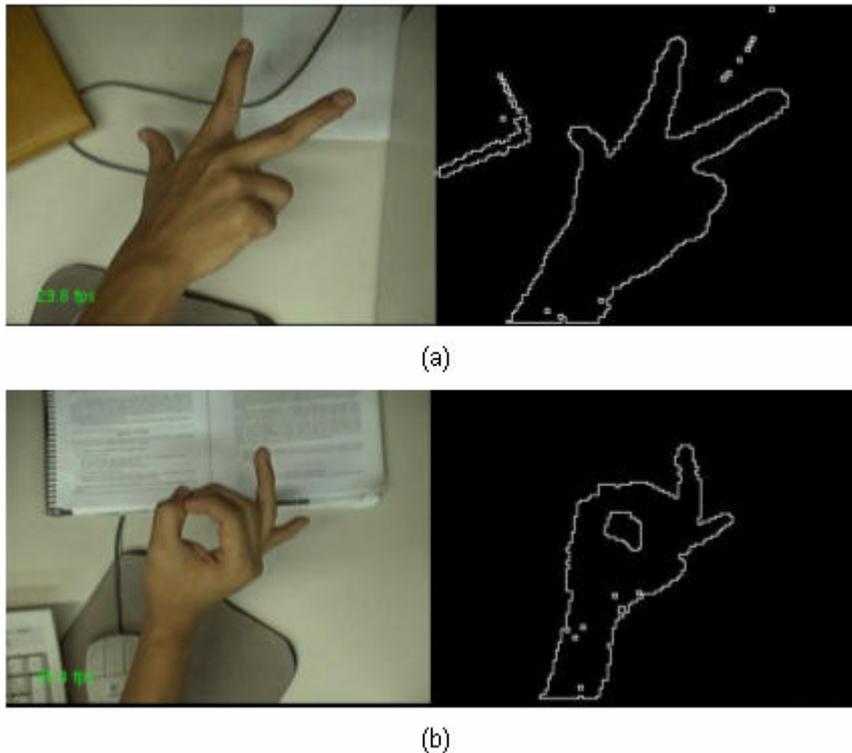


Figura 20 - Alguns resultados da detecção de contornos realizada em diferentes orientações da mão, (a) presença de ruídos, (b) presença de contornos internos ou buracos. 45 qps.

4.4. Detecção de Contornos

O problema de detecção da silhueta está fortemente relacionado como o processo de reconhecimento em si. Ainda que de certo modo o processo de detecção da silhueta possa ser visto como sendo um processo de reconhecimento (porque já identifica qual contorno representa um objeto), ele é considerado como sendo apenas um processo de identificação, diferente do processo de reconhecimento total, em que relações, características especiais e mudanças na morfologia dos objetos são reconhecidas e exploradas nas aplicações.

Para localizar a silhueta dos objetos, os contornos presentes na imagem são percorridos em algum sentido e em seguida são armazenados. Como as bordas são todas conexas os objetos estarão sempre representados por um ou mais contornos. Após o percurso e armazenamento dos contornos é feito um processo de seleção, no qual alguns contornos são descartados e outros são identificados como sendo os objetos. Assim existem duas etapas: uma que

percorre e armazena os objetos, e a outra de seleção, esta última mais diretamente ligada ao reconhecimento.

O percurso e armazenamento dos contornos é tradicionalmente feito numa ordem predeterminada, horária ou anti-horária. Partindo de um ponto qualquer do contorno os pontos são visitados seguindo a ordem estabelecida e o percurso é armazenado de forma que o contorno seja fielmente representado.

Lista de Códigos (*Chain Codes*) é uma das abordagens mais utilizadas para a representação dos contornos [40]. Como seu nome bem o diz, é uma lista de códigos onde, para cada ponto na lista, existe uma direção a seguir para achar o próximo ponto do contorno. As direções definidas nesta representação, na ordem anti-horária, podem ser vistas na Figura 21.

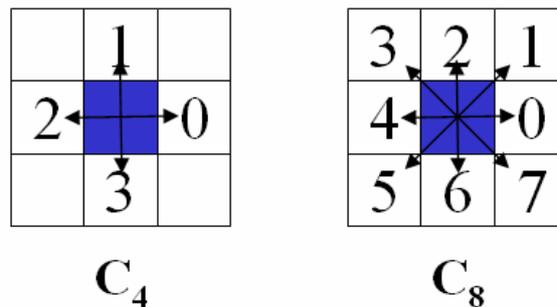


Figura 21 - Lista de Códigos (**Chain Codes**). Para cada ponto de referência (em azul) existem 4 ou 8 direções possíveis que se podem seguir na representação do contorno.

A Lista de Códigos é uma representação que traz algumas vantagens, principalmente no que se refere ao armazenamento. Considerando a vizinhança 8 (C_8) necessitamos apenas de 3 bits para representar um ponto na cadeia.

Definida a forma de representação dos pontos do contorno, o algoritmo parte de um ponto arbitrário escolhendo, em cada passo, um novo vizinho. As direções consideradas para a busca deste vizinho na nossa implementação foram as da vizinhança 8 (C_8) mostrada na Figura 21.

A idéia para percorrer os contornos é mostrada na Figura 22, em três passos. Nesta figura apenas é mostrado um trecho de um contorno maior, que serve para esquematizar alguns passos do algoritmo e entender a idéia de como é feito o percurso. Os pontos em branco representam o fundo e os pontos em preto representam o contorno que está sendo percorrido. Os pontos marcados com um “X” já fazem parte da silhueta, e os pontos onde as setas começam são

os que estão sendo analisados; estes também pertencem à silhueta. As setas que apontam para os pontos marcados com “X” indicam a direção do ponto que o antecede na silhueta, e as setas que apontam para os pontos marcados com “P” indicam o próximo ponto da silhueta. As setas curvas indicam a direção de busca (anti-horária) para procurar o próximo ponto da silhueta.

Todos os pontos de análise (onde começam as setas), a cada passo, têm um apontador para a direção do ponto que os antecede na silhueta. Tomando como referência essa direção e utilizando o sentido de busca (anti-horário), procura-se o próximo ponto que fará parte da silhueta; esse ponto é aquele que ainda não foi visitado, faz parte do contorno e é o primeiro a ser escolhido ou visitado a partir da direção inicial de busca. Em cada um dos passos (a), (b) e (c) as setas para os pontos “X” indicam a direção inicial de busca, as setas curvas indicam o sentido da busca e as setas para os pontos “P” apontam para o próximo ponto que faz parte da silhueta. O percurso do algoritmo tem dois critérios de parada: no caso de contornos fechados o algoritmo pára quando são alcançadas posições próximas do ponto inicial de partida, e no caso de contornos não fechados o algoritmo para quando não existem mais posições de busca na direção oposta à direção de partida.

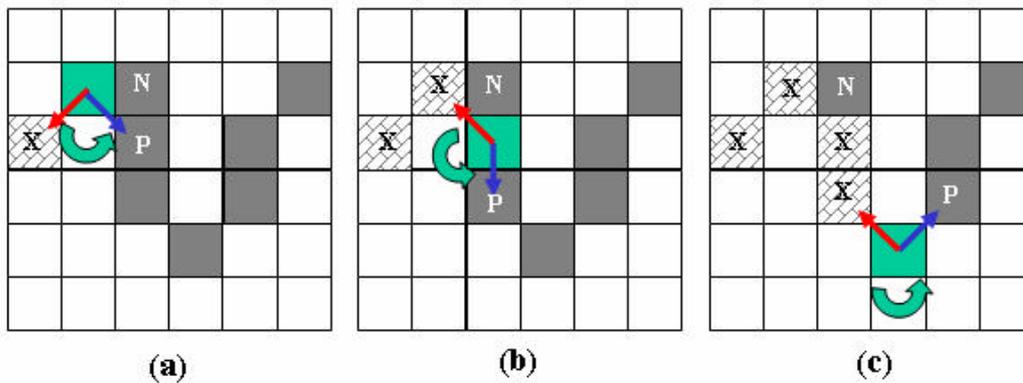


Figura 22 - Passos do algoritmo de percurso de contornos. A cada passo, novos pontos da silhueta são detectados e alguns pontos “N” são descartados. As setas curvas indicam o sentido de busca, as setas que apontam para os pontos marcados com “X” indicam a direção inicial de busca, as setas que apontam para os pontos marcados com “P” indicam o próximo ponto.

Como se pode ver no passo (a) da Figura 22, nem todos os pontos do contorno são considerados para fazer parte da silhueta, estes pontos não considerados estão marcados com “N”. Levando em conta que os contornos nem sempre são de largura unitária, muitos pontos são descartados ao longo do

percurso e assim pode-se achar um número menor ainda de pontos que definem a silhueta.

Na Figura 23 mostram-se alguns trechos de contornos com algumas particularidades produzidas durante o processo de detecção de silhueta. Nesta figura pode-se ver, dentro do contorno, um pequeno ciclo que deve ser considerado pelo algoritmo de percurso, já que este representa um problema para a robustez do algoritmo. Como se vê em (b) o percurso pode acabar prematuramente, afetando a detecção. Uma solução para este problema consiste em considerar o ponto “Z” como sendo uma ponte especial, e deixar a Lista de Códigos passar duas vezes por este ponto. O ponto “Z” é um ponto no meio do contorno (que não é nem começo e nem fim) para onde o percurso volta em uma direção diferente à da direção original de saída (utilizando uma aresta diferente). O percurso não permite se percorrer uma determinada aresta mais de uma vez.

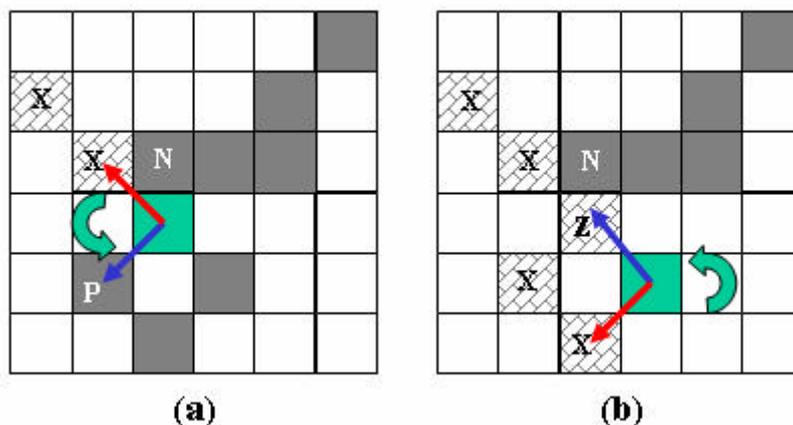


Figura 23 - Pequenos ciclos dentro dos contornos. Estes ciclos representam problemas para o algoritmo de percurso, pois como é indicado em (b) a detecção da silhueta pode acabar muito antes de percorrer o contorno inteiro, o que traz problemas na detecção.

Uma vez que se tenham percorrido todos os contornos da imagem, as silhuetas dos objetos são armazenadas em um conjunto de listas. O próximo passo é a seleção das listas de códigos que representam o objeto de interesse.

4.5. Seleção da Silhueta da Mão

No processo de seleção existe um problema de decisão. No caso de objetos ou gestos da mão que são representados por silhuetas internas e externas, qual ou quais silhuetas considerar? Considerar apenas a silhueta

externa? Ou considerar todas as silhuetas que o objeto apresenta? A alternativa escolhida para lidar com este problema afeta diretamente a abordagem utilizada no reconhecimento, trazendo as suas vantagens e desvantagens.

No caso particular dos gestos da mão, sabe-se que um dos objetos mais representativos na imagem é a própria mão. O tamanho da silhueta externa da mão é muito maior do que as silhuetas internas e do que os ruídos presentes nas imagens. Com base nesta hipótese a detecção da silhueta da mão é feita considerando, em número de pontos, a maior silhueta encontrada na imagem.

Para expandirmos nossa aplicação para reconhecer duas mãos, necessitaremos de uma abordagem diferente e mais complexa. A abordagem utilizada neste trabalho é baseada apenas na silhueta externa de uma única mão.

A Figura 24 ilustra os resultados da aplicação de detecção de silhueta nas imagens das Figuras 19 e 20 utilizando o critério de só considerarmos a maior Lista de Códigos.

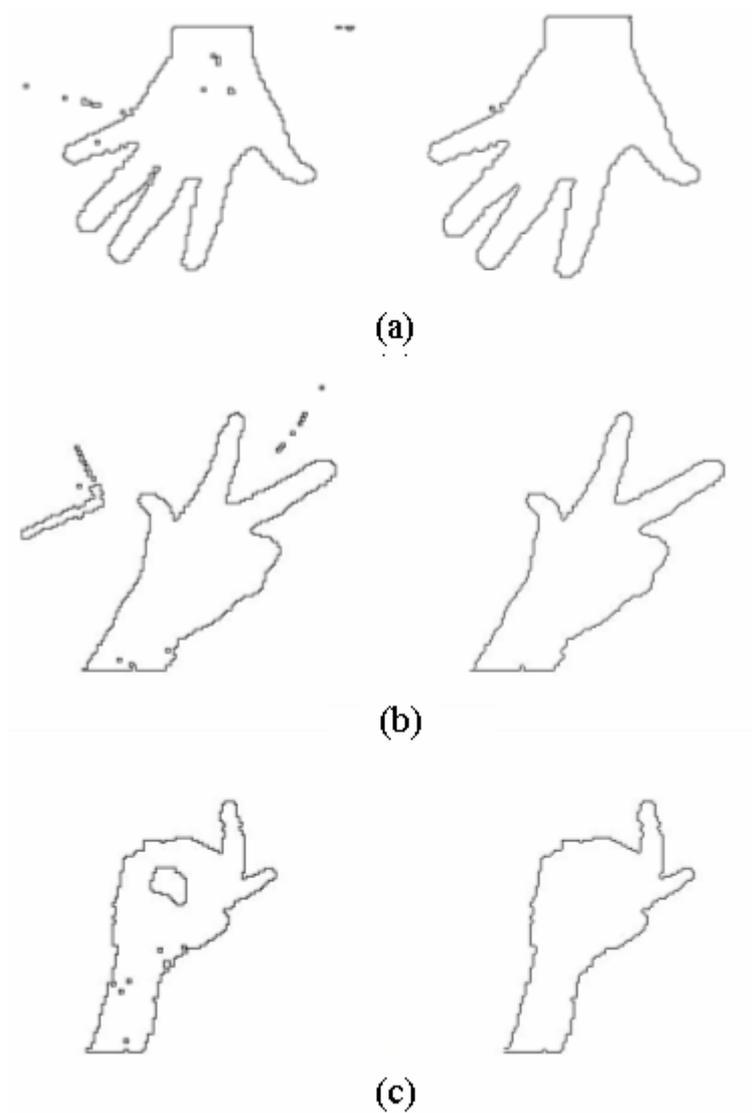


Figura 24 - Resultados do processo de detecção da silhueta da mão aplicados à imagem da Figura 19 em (a), e às imagens da Figura 20 em (b) e (c). Em (a) e (b) são removidos os ruídos e em (c) não é considerada a silhueta interna. Número médio de quadros por segundo igual a 35 qps.

As imagens da Figura 24 representam o resultado final da etapa de detecção de silhueta. A amostragem dos pontos da silhueta pode ser mudada para se ter uma amostragem menor, descartando assim um número maior de pontos da silhueta. Esta subamostragem pode ser feita sem afetar a morfologia da mão até um certo limite. A Figura 25 mostra a detecção de uma silhueta com uma subamostragem contendo um quinto dos pontos originalmente detectados.



Figura 25 - Subamostragem dos pontos na silhueta detectada. A imagem à direita tem um quinto dos pontos da imagem à esquerda.

4.6. Considerações Finais do Capítulo

A abordagem utilizada é baseada na silhueta externa da mão; uma única silhueta é considerada e a etapa de identificação de características no reconhecimento é feita sobre o conjunto de pontos que representa a silhueta da mão.

A decisão de levar em conta as silhuetas externas certamente restringe e ao mesmo tempo facilita os processos que serão realizados na etapa do reconhecimento. É possível utilizar somente silhuetas externas em uma grande quantidade de aplicações, mas pode-se realizar tarefas de reconhecimento mais complexas utilizando mais informações provenientes do processo de detecção de silhueta.